

CLAIMS

What is claimed is:

1. A method of correlating gene expression with genetic variations comprising:
  - obtaining a first plurality of gene expression profiles from a plurality of individuals with a first genotype;
  - obtaining a second plurality of gene expression profiles from a plurality of individuals with a second genotype;
  - comparing the first and second gene expression profiles; and
  - indicating the genes whose expression segregates with the genotypes as the genes affected by the genotypes.
2. The method of Claim 1 wherein the genotypes are the states of a SNP.
3. The method of Claim 1 wherein the genotypes are haplotypes.
4. The method of Claim 2 wherein the gene expression profiles have at least 2 genes.
5. The method of Claim 4 wherein the gene expression profiles have at least 500 genes.
6. The method of Claim 5 wherein the gene expression profiles have at least 1000 genes.
7. The method of Claim 6 wherein the gene expression profiles have at least 5000 genes.
8. The method of Claim 7 wherein the gene expression profiles have at least 10000 genes.

9. The method of Claim 2 wherein the step of comparing comprises a step of evaluating the difference in gene expression between the first and second genotypes.

5

10. The method of Claim 9 wherein the step of evaluating comprises calculating a normalized difference in gene expression between the first and second genotypes.
11. The method of Claim 10 wherein the step of comparing comprises a step of calculating a SNPmetric for each SNP and each gene according to:

$$^{gene} \Gamma_{SNP} = \frac{(E_{wt}^{avg} - E_{e/o}^{avg})^c}{\sigma_{wt}^a \sigma_{e/o}^b}$$

wherein:  $^{gene} \Gamma_{SNP}$  = SNPmetric for a given gene;

$E_{wt}^{avg}$  = average gene expression for wild type SNP for the gene;

$E_{e/o}^{avg}$  = average gene expression for heterozygous/homozygous mutant for the gene;

$\sigma_{wt}$  = standard deviation of gene expression of wild type SNP for the gene;

$\sigma_{e/o}$  = standard deviation of gene expression of heterozygous/homozygous mutant for the gene; and

$a, b, c$  = sensitivity parameters.

20

25

12. The method of Claim 10 wherein the step of comparing comprises a step of calculating a SNPmetric for each SNP and each gene according to:

$$^{gene} \Gamma_{SNP} = \frac{|(E_{wt}^{avg} - E_{e/o}^{avg})|^c}{\sigma_{wt}^a \sigma_{e/o}^b}$$

wherein:  $\Gamma_{SNP}^{gene}$  = SNPmetric for a given gene;  
 $E_{wt}^{avg}$  = average gene expression for wild type SNP for the gene;  
 $E_{e/o}^{avg}$  = average gene expression for heterozygous/homozygous

5 mutant for the gene;

$\sigma_{wt}$  = standard deviation of gene expression of wild type SNP for  
the gene;

$\sigma_{e/o}$  = standard deviation of gene expression of  
heterozygous/homozygous mutant for the gene; and

10  $a, b, c$  = sensitivity parameters.

13. A computer software product for correlating gene expression with genetic variations comprising:

computer program code that inputs a first plurality of gene expression profiles from a plurality of individuals with a first genotype;

computer program code that inputs a second plurality of gene expression profiles from a plurality of individuals with a second genotype;

computer program code that compares the first and second gene expression profiles;

computer program code that indicates the genes whose expression segregates with the genotypes as the genes affected by the genotypes; and

a computer readable medium for storing the codes.

14. The computer software product of Claim 13 wherein the genotypes are the

25 states of a SNP.

15. The computer software product of Claim 13 wherein the genotypes are haplotypes.

30 16. The computer software product of Claim 13 wherein the gene expression profiles have at least 2 genes.

17. The computer software product of Claim 16 wherein the gene expression profiles have at least 500 genes.
- 5 18. The computer software product of Claim 17 wherein the gene expression profiles have at least 1000 genes.
- 10 19. The computer software product of Claim 18 wherein the gene expression profiles have at least 5000 genes.
20. The computer software product of Claim 19 wherein the gene expression profiles have at least 10,000 genes.
21. The computer software product of Claim 14 wherein the code that compares comprises code that evaluates the difference in gene expression between the first and second genotypes.
22. The computer software product of Claim 21 wherein the code that evaluates comprises code that calculates a normalized difference in gene expression between the first and second genotypes.
23. The computer software product of Claim 22 wherein the code that compares comprises code that calculates a SNPmetric for each SNP and

each gene according to:

$$\text{gene } \Gamma_{SNP} = \frac{(E_{wt}^{avg} - E_{e/o}^{avg})^c}{\sigma_{wt}^a \sigma_{e/o}^b}$$

25 wherein:  $\text{gene } \Gamma_{SNP}$  = SNPmetric for a given gene;  
 $E_{wt}^{avg}$  = average gene expression for wild type SNP for the gene;  
 $E_{e/o}^{avg}$  = average gene expression for heterozygous/homozygous mutant for the gene;

$\sigma_{wt}$  = standard deviation of gene expression of wild type SNP for the gene;

$\sigma_{e/o}$  = standard deviation of gene expression of heterozygous/homozygous mutant for the gene; and

5                    $a, b, c$  = sensitivity parameters.

24. The computer software product of Claim 22 wherein the code that compares comprises code that calculates a SNPmetric for each SNP and

each gene according to:

$$\Gamma_{SNP}^{gene} = \frac{|(E_{wt}^{avg} - E_{e/o}^{avg})|^c}{\sigma_{wt}^a \sigma_{e/o}^b}$$

10                   wherein:      $\Gamma_{SNP}^{gene}$  = SNPmetric for a given gene;  
 $E_{wt}^{avg}$  = average gene expression for wild type SNP for the gene;  
 $E_{e/o}^{avg}$  = average gene expression for heterozygous/homozygous mutant for the gene;

15                    $\sigma_{wt}$  = standard deviation of gene expression of wild type SNP for the gene;  
 $\sigma_{e/o}$  = standard deviation of gene expression of heterozygous/homozygous mutant for the gene; and  
 $a, b, c$  = sensitivity parameters.

- 20                   25. A computer readable medium comprising computer-executable instructions for performing the method for correlating gene expression with genetic variations comprising:  
                      inputting a first plurality of gene expression profiles from a plurality of individuals with a first genotype;  
                      inputting a second plurality of gene expression profiles from a plurality of individuals with a second genotype;  
                      comparing the first and second gene expression profiles; and

indicating the genes whose expression segregates with the genotypes as the genes affected by the genotypes.

- 5            26. The computer readable medium of Claim 25 wherein the genotypes are the states of a SNP.
- 10            27. The computer readable medium of Claim 25 wherein the genotypes are haplotypes.
- 15            28. The computer readable medium of Claim 25 wherein the gene expression profiles have at least 2 genes.
- 20            29. The computer readable medium of Claim 28 wherein the gene expression profiles have at least 500 genes.
- 25            30. The computer readable medium of Claim 29 wherein the gene expression profiles have at least 1000 genes.
- 30            31. The computer readable medium of Claim 30 wherein the gene expression profiles have at least 5000 genes.
- 35            32. The computer readable medium of Claim 31 wherein the gene expression profiles have at least 10000 genes.
- 40            33. The computer readable medium of Claim 32 wherein the step of comparing comprises a step of evaluating the difference in gene expression between the first and second genotypes.
- 45            34. The computer readable medium of Claim 26 wherein the step of evaluating comprises calculating a normalized difference in gene expression between the first and second genotypes.

35. The computer readable medium of Claim 34 wherein the step of comparing comprises a step of calculating a SNPmetric for each SNP and

$$\text{each gene according to: } {}^{gene} \Gamma_{SNP} = \frac{(E_{wt}^{avg} - E_{e/o}^{avg})^c}{\sigma_{wt}^a \sigma_{e/o}^b}$$

5

wherein:  $\Gamma_{SNP}$  = SNPmetric for a given gene;  
 $E_{wt}^{avg}$  = average gene expression for wild type SNP for the gene;  
 $E_{e/o}^{avg}$  = average gene expression for heterozygous/homozygous mutant for the gene;

$\sigma_{wt}$  = standard deviation of gene expression of wild type SNP for the gene;  
 $\sigma_{e/o}$  = standard deviation of gene expression of heterozygous/homozygous mutant for the gene; and  
 $a, b, c$  = sensitivity parameters.

- 10 36. The computer readable medium of Claim 34 wherein the step of comparing comprises a step of calculating a SNPmetric for each SNP and

$$\text{each gene according to: } {}^{gene} \Gamma_{SNP} = \frac{|(E_{wt}^{avg} - E_{e/o}^{avg})|^c}{\sigma_{wt}^a \sigma_{e/o}^b}$$

15

20 wherein:  $\Gamma_{SNP}$  = SNPmetric for a given gene;  
 $E_{wt}^{avg}$  = average gene expression for wild type SNP for the gene;  
 $E_{e/o}^{avg}$  = average gene expression for heterozygous/homozygous mutant for the gene;

25  $\sigma_{wt}$  = standard deviation of gene expression of wild type SNP for the gene;  
 $\sigma_{e/o}$  = standard deviation of gene expression of heterozygous/homozygous mutant for the gene; and

*a, b, c* =sensitivity parameters.

37. A system for associating a genotype with gene expression comprising:  
a processor; and

5  
a memory coupled with the least one processor, the memory storing a plurality of machine instructions that cause the processor to perform logical steps, wherein the logical steps include:

obtaining a first plurality of gene expression profiles from a plurality of individuals with a first genotype;

10  
obtaining a second plurality of gene expression profiles from a plurality of individuals with a second genotype;

comparing the first and second gene expression profiles; and

indicating the genes whose expression segregates with the genotypes as the genes affected by the genotypes.

38. The system of Claim 37 wherein the genotypes are the states of a SNP.

39. The system of Claim 37 wherein the genotypes are haplotypes.

40. The system of Claim 37 wherein the gene expression profiles have at least 2 genes.

41. The system of Claim 40 wherein the gene expression profiles have at least 500 genes.

42. The system of Claim 41 wherein the gene expression profiles have at least 1000 genes.

43. The system of Claim 42 wherein the gene expression profiles have at least 5000 genes.

44. The system of Claim 43 wherein the gene expression profiles have at least 10000 genes.
- 5 45. The system of Claim 38 wherein the step of comparing comprises a step of evaluating the difference in gene expression between the first and second genotypes.
- 10 46. The system of Claim 45 wherein the step of evaluating comprises calculating a normalized difference in gene expression between the first and second genotypes.
47. The system of Claim 46 wherein the step of comparing comprises a step of calculating a SNPmetric for each SNP and each gene according to:

$$\text{gene } \Gamma_{SNP} = \frac{(E_{wt}^{avg} - E_{e/o}^{avg})^c}{\sigma_{wt}^a \sigma_{e/o}^b}$$

wherein:  $\Gamma_{SNP}$  = SNPmetric for a given gene;  
 $E_{wt}^{avg}$  = average gene expression for wild type SNP for the gene;  
 $E_{e/o}^{avg}$  = average gene expression for heterozygous/homozygous

mutant for the gene;  
20  $\sigma_{wt}$  = standard deviation of gene expression of wild type SNP for the gene;

$\sigma_{e/o}$  = standard deviation of gene expression of heterozygous/homozygous mutant for the gene; and  
a, b, c = sensitivity parameters.

48. The system of Claim 46 wherein the step of comparing comprises a step of calculating a SNPmetric for each SNP and each gene according to:

$$^{gene} \Gamma_{SNP} = \frac{|(E_{wt}^{avg} - E_{e/o}^{avg})|^c}{\sigma_{wt}^a \sigma_{e/o}^b}$$

5           wherein:      $^{gene} \Gamma_{SNP}$  = SNPmetric for a given gene;  
                     $E_{wt}^{avg}$  = average gene expression for wild type SNP for the gene;  
                     $E_{e/o}^{avg}$  = average gene expression for heterozygous/homozygous  
mutant for the gene;  
                     $\sigma_{wt}$  = standard deviation of gene expression of wild type SNP for  
the gene;  
                     $\sigma_{e/o}$  = standard deviation of gene expression of  
heterozygous/homozygous mutant for the gene; and  
                     $a, b, c$  = sensitivity parameters.